European Journal of Clinical Nutrition (2006) 60, 971–977
 0954-3007/06 $30.00
www.nature.com/ejcn

ORIGINAL ARTICLE

Validity and reliability of a new food frequency questionnaire compared to 24 h recalls and biochemical measurements: pilot phase of Golestan cohort study of esophageal cancer

AF Malekshah[1,4], M Kimiagar[1,4], M Saadatian-Elahi[2,3], A Pourshams[1], M Nouraie[1], G Goglani[1], A Hoshiarrad[4], M Sadatsafavi[1], B Golestan[4], A Yoonesi[1], N Rakhshani[1], S Fahimi[1,2], D Nasrollahzadeh[1], R Salahi[5], A Ghafarpour[4], S Semnani[5], JP Steghens[3], CC Abnet[6], F Kamangar[6], SM Dawsey[6], P Brennan[2], P Boffetta[2] and R Malekzadeh[1]

[1]*Digestive Disease Research Center, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran;* [2]*International Agency for Research on Cancer, Lyon, France;* [3]*Edouard Herriot Hospital and EA 3090, Lyon, France;* [4]*National Nutrition and Food Science Research Institute of Shaheed Beheshti University of Medical Science, Tehran, Iran;* [5]*Golestan University of Medical Sciences, Gorgan, Iran and* [6]*Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA*

Background: A pilot study was carried out to evaluate validity and reproducibility of a food frequency questionnaire (FFQ), which was designed to be used in a prospective cohort study in a population at high risk for esophageal cancer in northern Iran.
Methods: The FFQ was administered four times to 131 subjects, aged 35–65 years, of both sexes. Twelve 24-h dietary recalls for two consecutive days were administered monthly during 1 year and used as a reference method. The excretion of nitrogen was measured on four 24-h urine samples, and plasma levels of β-carotene, retinol, vitamin C and α-tocopherol was measured from two time points. Relative validity of FFQ and 24-h diet recall was assessed by comparing nutrient intake derived from both methods with the urinary nitrogen and plasma levels of β-carotene, retinol, vitamin C and α-tocopherol.
Results: Correlation coefficients comparing energy and nutrients intake based on the mean of the four FFQ and the mean of twelve 24-h diet recalls were 0.75 for total energy, 0.75 for carbohydrates, 0.76 for proteins and 0.65 for fat. Correlation coefficients between the FFQ-based intake and serum levels of β-carotene, retinol, vitamin C and vitamin E/α-tocopherol were 0.37, 0.32, 0.35 and 0.06, respectively. Correlation coefficients between urinary nitrogen and FFQ-based protein intake ranged from 0.23 to 0.35. Intraclass correlation coefficients used to measure reproducibility of FFQ ranged from 0.66 to 0.89.
Conclusion: We found that the FFQ provides valid and reliable measurements of habitual intake for energy and most of the nutrients studied.
European Journal of Clinical Nutrition (2006) **60,** 971–977. doi:10.1038/sj.ejcn.1602407; published online 8 February 2006

Keywords: validity; reliability; esophageal cancer; cohort study; Iran

Correspondence: Professor R Malekzadeh, Digestive Disease Research Center, Tehran University Medical Sciences, Tehran, Iran.
E-mail: malek@ams.ac.ir
Guarantor: R Malekzadeh.
Contributors: The authors have contributed to the manuscript by planning the study (DDRC, IARC, NCI, NNFRI), collecting data (DDRC, NNFRI), calculating the nutritional components (DDRC, NNFRI), serum vitamin measurements (Edouard Herriot Hospital, Addenbrooke's Hospital in Cambridge,UK), analysis of data (DDRC, IARC,NCI), and preparation and revision of the manuscript (all authors). DDRC: Digestive Disease Research Center, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran; IARC: International Agency for Research on Cancer, Lyon, France; NNFR: National Nutrition and Food Science Research Institute of Shaheed Beheshti University of Medical Science, Iran; NCI: Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics National Cancer Institute, Bethesda, MD, USA; Edouard Herriot Hospital and EA 3090, Lyon, France; Addenbrooke's Hospital, Unit of Clinical Biochemistry, Cambridge, UK.
Received 26 July 2005; revised 28 November 2005; accepted 21 December 2005; published online 8 February 2006

Introduction

In the 1970s, population-based cancer registration data established that the Eastern part of Mazandaran province (now Golestan Province, Iran) was an area with very high rates of oesophageal cancer (Kmet and Mahboubi, 1972). Etiologic factors responsible for these high rates are largely unknown (Islami *et al*., 2004), and only two studies have examined the role of diet in association with oesophageal cancer in this area (Hormozdiari *et al*., 1975; Cook-Mozaffari *et al*., 1979). Residents in the high incidence areas had lower intake of vitamins A and C, riboflavin, animal proteins and fresh vegetables and fruits and a higher consumption of sheep and goat milk than that of residents in other areas (Hormozdiari *et al*., 1975).

In order to further investigate the causes of the high rates of oesophageal cancer in Golestan, including novel dietary hypotheses, an initiative involving the Digestive Diseases Research Center (DDRC) of Tehran University of Medical Sciences, the International Agency for Research on Cancer (IARC) and the US National Cancer Institute (NCI) has been established to conduct a large-scale, prospective cohort study in this region. To assess the feasibility of such a cohort study, a pilot study was recently completed (Pourshams *et al*., 2005).

As part of this cohort study, we plan to administer a new food frequency questionnaire (FFQ) designed specifically to capture the dietary practices of the study participants. To evaluate the validity and reliability of this FFQ, we conducted a study over 12 months during which subjects were administered four FFQ and twelve 24-h dietary recalls and had serum and urine collected at two time points to measure biomarkers of food intake. Among biomarkers of dietary intake, the 24-h urinary excretion of nitrogen and the plasma concentration of carotenoids are used widely in epidemiological studies to estimate the average 24-h protein intake (Ocke *et al*., 1997b; Prentice *et al*., 2002; Bingham, 2003) and the habitual intake of fruits and vegetables (Bohlscheid-Thomas *et al*., 1997a; Kaaks *et al*., 1997; Ocke *et al*., 1997a), respectively. Although the first one is a marker of absolute quantitative intake of protein (Kaaks *et al*., 1997), the second one provides only an estimation of intake levels. Plasma vitamin C has also been used to reflect recent intake of vitamin C.

This paper describes the reliability and validity of the FFQ developed for the cohort study planned for Golestan province.

Materials and methods

Subjects

Study subjects were participants in the pilot phase of a cohort study on esophageal cancer in the Golestan province in northeastern Iran (Pourshams *et al*., 2005). Between June 2002 and June 2003, a dietary survey was conducted on a subsample of 142 subjects selected by stratified random sampling regarding living place (rural, urban) from the 1349 subjects recruited in the pilot phase. They comprised 57 subjects from Gonbad, the largest city in the study area, and 85 subjects from three surrounding villages (Incheborun, Hali-Akhond and Aq-Abad). The dietary survey included an FFQ and a 24-h recall. Each subject was investigated for 12 months during which four FFQs, twelve 24-h dietary recalls, four 24-h urine samples and two blood samples were collected. The study was not conducted during the month of Ramadan (September 2003) when the Muslims refrain from foods and drinks from dawn to dusk (Table 1).

The study protocol and the informed consent used for this investigation were developed, implemented and approved by the ethical review boards of DDRC and the IARC. The analysis of data and samples was approved by ethical review committee of the NCI.

Dietary questionnaires

A semi-quantitative FFQ (Boeing *et al*., 1997; Bohlscheid-Thomas *et al*., 1997a, b; Kroke *et al*., 1999a) was designed and developed to measure the intake of common food groups, energy and nutrients over the previous year. As the initial step, an FFQ draft including 158 single food items was designed based on the information provided by experienced nutritionists familiar with the local diet. Spaces were included to capture information on additional and miscellaneous items. Frequency of food intake was recorded in times per day, week, month and year or as never. This first draft was administered to 142 subjects in the month before the beginning of the pilot study. Items that were reported to have been consumed rarely or never were excluded from the final version of the FFQ. We enquired about individual food items, rather than recipes, because of individual variations in recipe preparation. The number of food items was reduced to 150 and portion size was estimated as the median of each item. Pictures of different portion sizes were used for 51 food

Table 1 Study timeline for the administration of the FFQs, 24-h dietary recalls, blood samples and urine samples

Month	*1*	*2*	*3*	*#*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
FFQ			×			×			×			×	
24-h recall	×	×	×		×	×	×	×	×	×	×	×	×
Blood sample			×						×				
Urine sample			×			×			×			×	

#Ramadan month.
Abbreviations: FFQ, Food frequency questionnaire.

items to increase the precisions of the estimates (EPIC Group of Spain, 1997c). The FFQ was completed at the subjects' home and then reviewed to identify incomplete and illogical response. For the estimate of nutrients, recipes and complex foods were broken down into simple foods. For example, fried, battered or crumbed foods were broken down into simple foods plus the corresponding fat.

Twelve 24-h dietary recalls were completed once a month over a period of a year for all 131 subjects, except for the month of Ramadan. All food and beverage items consumed the previous day from the time of waking up to going to bed at night were recorded on an open questionnaire. All days of the week, except weekends were included in the 24-h recalls. The 24-h dietary recalls were administered in face-to-face interviews at the subjects' home and in their native language by four trained interviewers and were then checked by a nutritionist for completeness. A Soehnle scale (± 5 g) was used to improve the precision of the estimation of dietary intake (Lucas *et al*., 1995). The interview lasted about 40–50 min on the first day and 20–30 min on the second day. Participants were interviewed by different personnel at different sessions to reduce interviewer bias.

Biological sample collection and biochemical measurements

Four 24-h urine samples were collected quarterly to determine the amount of daily urinary nitrogen in grams. Subjects were provided a container prefilled with 10 g boric acid to collect urine over the next 24-h. Urine samples were considered incomplete when less or more than normal range of creatinine (11–20 mg/kg/day for women and 14–26 mg/kg/day for men) was present as measured by colorimetric Jaffe method (Hristova and Henry, 2001). Urea nitrogen excretion was measured by spectrophotometry, and converted to urea nitrogen excretion in g/day (Hristova and Henry, 2001). We assumed that urea nitrogen excretion was a constant proportion (85%) of total urinary nitrogen, so total protein intake was derived from the formula 6.25 (urinary nitrogen + 2), as suggested by Isaksson (1980) and Kroke *et al*. (1999a).

During the third and eighth month of the project, from each subject, blood samples (8 ml) were collected in tubes without additives or anticoagulants. Immediately, samples were protected from light by using aluminum foil, kept in a cool box (-8 to $+1^{\circ}$C) for a maximum of 2 h. To allow for clot retraction, samples were centrifuged at 3000 r.p.m. for 10 min. Serum was extracted and 5% meta-phosphoric acid was used to preserve sample for vitamin C measurement. In total, five tubes with 2 ml of serum per subject were stored at -80°C.

The analysis of vitamin C was performed using a fluorometric assay (Vuilleumier and Keck, 1989) at the unit of Clinical Biochemistry of Addenbrooke's Hospital in Cambridge, UK. A reversed-phase high-pressure liquid chromatography procedure was used for the quantitative measurements, in serum, of seven carotenoids (lutein, zeaxanthin, canthaxanthin, β-cryptoxanthin, lycopene, α-carotene and β-carotene), retinol and α-tocopherol (Steghens *et al*., 1997) at the Laboratory of Biochemistry of the Edouard Herriot Hospital, in Lyon, France. After extraction with hexane, all compounds were separated on an Adsorbosphere HS C18 column and analyzed at four different wavelengths (292, 325, 450 and 473 nm). Total cholesterol was measured by the enzymatic method (Bachor *et al*., 2001) by auto analyzer instrument (Hitachi 704, Alcyon300) in the biochemistry laboratory of the Shariati Hospital, Tehran, Iran.

Analysis and statistical methods

For all main food items in the FFQ, the frequency per day was multiplied by the amount consumed, based on portion size, to compute the total amount consumed per day. The food items were then grouped into 14 categories.

Nutritionist software version IV (Nutritionist IV, Version 3.5.2) was used to calculate daily energy and nutrient intake. The USDA food composition table (FCT) was used for most items (USDA, Release 11, 1994). For some items such as bread, vetch, pepper green, wild plum, mint, sweet canned cherry and sour cherry, the Iranian food composition table was consulted (Azar and Sarkisian, 1981). The normality of the dietary intake variable distributions was assessed by Q–Q plots and found to be skewed. Log transformation improved normality and these values were used throughout the analysis. Descriptive analyses were performed using mean and s.d. Crude Pearson's correlation coefficients were used to measure the agreement between nutrient intakes from the FFQ and both the 24-h diet recalls and biomarkers measured in urine and blood. Partial correlation coefficients adjusted for serum cholesterol were calculated for vitamin E concentration (Hunter, 1990).

The reproducibility of FFQs was examined by intra-class correlation coefficients, using a two-way random model between nutrient intakes and FFQs. All analyses were performed for all subjects, and for men and women separately. Statistical analysis was performed with SPSS-software, version 11.5 (SPSS, Version 11.5.0, SPSS Inc. Chicago).

Results

Of the 142 subjects invited to participate, 10 were excluded because of incomplete data on an FFQ or 24-h recall, and one was excluded because of adherence to a special dietary regimen. The analysis is based on 131 individuals including 51 men and 80 women. Selected demographic characteristics of the study population are presented in Table 2. Compared to women, men were slightly older and had lower body mass index. Prevalence of current smoker in men and women was 22.4 and 1.2%, respectively.

The mean values of the daily intake of nutrients, based on the four FFQs and the twelve 24-h diet recalls are reported in Table 3. The FFQ tended to overestimate intake compared to

24-h recalls, especially in women. These overestimations were particularly notable for the vitamins.

Table 4 shows the correlation coefficients between the mean intake from the 24-h recalls and each FFQ plus the mean of the four FFQs. Correlation coefficients were higher while using the mean of the four FFQ values than using each FFQ separately. Correlation coefficients were high for cholesterol, but were lower for vitamin E, poly unsaturated fatty acid and mono unsaturated fatty acid (MUFA). Correlation coefficients were generally higher in women than in men, except for fat and its components (data not shown).

Serum concentrations of the measured biochemical markers by time of blood collection are presented in Table 5. Concentrations of all biomarkers, except for α-carotene were significantly higher in the second than in the first measurement. The blood levels of vitamin C and carotenoids were not different between smokers and non-smokers.

The correlation coefficients between blood biochemical markers and nutrient dietary intakes assessed by the FFQ and the 24-h diet recalls are reported in Table 6. In general, the correlation coefficients for FFQ3 were higher than for the other FFQs and similar to those observed for the 24-h diet recalls. Table 6 shows that this may not be an artefact except for α-tocopherol that is highly correlated with the same time blood sample. All correlation coefficients were higher for men than for women, except retinol (data not shown).

Table 2 Demographic characteristics of study subjects by sex

Variable	*Men (N = 49)*	*Women (N = 82)*
Age (year)	51.2 (13.2)	49.9 (9.8)
Height (cm)	167.7 (6.8)	153.5 (5.1)
Weight (kg)	67.6 (12.1)	62.3 (13.8)
BMI (kg/m^2)	24.0 (3.9)	26.5 (5.8)

Mean (standard deviation).
Abbreviations: BMI, body mass index.

The reproducibility of FFQ-based estimates of energy and nutrient intake are presented in Table 7. Intraclass correlation coefficients were equal to or higher than 0.75 for all nutrients, except for total fat and fat components. Intraclass

Table 4 Correlation coefficients of nutrient intake between the mean of twelve 24-h dietary recalls and each FFQ of the mean of all four FFQs

Nutrient	*FFQ1*	*FFQ2*	*FFQ3*	*FFQ4*	*Mean of FFQ*
Energy	0.58	0.60	0.56	0.62	0.75
Carbohydrate	0.58	0.53	0.51	0.66	0.75
Protein	0.57	0.49	0.57	0.73	0.76
Fat	0.44	0.58	0.56	0.44	0.65
SFA	0.56	0.63	0.68	0.61	0.75
MUFA	0.35	0.53	0.50	0.37	0.58
PUFA	0.34	0.48	0.40	0.25	0.52
Cholesterol	0.67	0.70	0.72	0.72	0.82
Vitamin C	0.59	0.51	0.65	0.52	0.65
Retinol	0.56	0.48	0.52	0.57	0.59
Vitamin E	0.27	0.45	0.44	0.40	0.49
β-Carotene	0.53	0.58	0.65	0.56	0.68

Abbreviations: FFQ, Food frequency questionnaire; MUFA, mono unsaturated fatty acid; PUFA, poly unsaturated fatty acid; SFA, saturated fatty acid.

Table 5 Serum concentrations of the biochemical markers according to the first and second time measurement

Biomarker (N)	*First measurement*	*Second measurement*	P*
Vitamin C (103)	19.21 (7.80)	25.68 (11.03)	0.001
Lutein (121)	0.31 (0.16)	0.35 (0.20)	0.03
Zeaxanthin (117)	0.053 (0.02)	0.068 (0.04)	0.001
Lycopene (120)	1.34 (0.75)	1.56 (0.64)	0.001
α-Carotene (65)	0.04 (0.02)	0.05 (0.03)	0.51
β-Carotene (121)	0.164 (0.12)	0.198 (0.15)	0.001
β-Cryptoxanthin (118)	0.04 (0.03)	0.13 (0.19)	0.001

Mean (s.d.) in μmol/l.
**P* for Paired samples t-test.
First measurement (month 3) and second measurement (month 8) refer to the two different times of serum collection.

Table 3 Daily intake of energy and nutrients, based on the mean of the twelve 24-h recalls and four FFQs

Nutrient	*Men*		*Women*	
	24-h recalls	*FFQs*	*24-h recalls*	*FFQs*
Energy (kcal)	2723 (555)	2696 (646)	1859 (450)	2101 (526)
Carbohydrates (g)	452 (104)	430 (111)	313 (78)	328 (96)
Protein (g)	73 (17)	67 (19)	48 (12)	51 (14)
Fat (g)	71 (18)	83 (25)	48 (17)	68 (25)
SFA (g)	16 (6)	18 (7)	10 (4)	15 (6)
MUFA (g)	27 (7)	30 (10)	18 (7)	24 (10)
PUFA (g)	19 (5)	23 (8)	13 (5)	19 (8)
Cholesterol (mg)	158 (94)	143 (102)	87 (55)	97 (65)
Vitamin C (mg)	45 (21)	89 (54)	37 (24)	69 (42)
Vitamin A (RE)	600 (638)	410 (300)	379 (381)	336 (264)
Vitamin E (mg)	0.95 (1.1)	0.65 (0.7)	0.68 (1.0)	0.66 (1.4)
β-Carotene (μg)	90 (107)	156 (184)	66 (78)	156 (21)

Mean (standard deviation).
Abbreviations: FFQ, Food frequency questionnaire; MUFA, mono unsaturated fatty acid; PUFA, poly unsaturated fatty acid; SFA, saturated fatty acid.

Table 6 Pearson's correlation coefficients between serum biomarkers and estimated intakes based on twelve 24 h diet recalls or four FFQs

Biomarker[a]	*Mean of twelve 24-h recalls*	*Mean of four FFQs*	*FFQ1 vs blood 1*	*FFQ1*[b]	*FFQ2*	*FFQ3 vs blood 2*	*FFQ3*[b]	*FFQ4*
Vitamin C	0.37	0.35	0.27	0.33	0.25	0.48	0.34	0.25
Retinol	0.26	0.32	0.17	0.26	0.29	0.25	0.27	0.27
β-Carotene	0.35	0.37	0.19	0.27	0.31	0.43	0.44	0.35
α-tocopherol[c]	0.10	0.06	0.19	−0.01	0.03	0.16	0.12	0.07
Urine nitrogen	0.38	0.37	0.27[d]	0.29	0.27	0.12[d]	0.23	0.35

[a]Average of two measurements at two time points and all markers in serum, except nitrogen.
[b]At the time of administration of this FFQ, one blood sample was taken.
[c]Adjusted for serum cholesterol and compared to vitamin E as estimated by questionnaire.
[d]Correlation between FFQ1 and Urine 1 and FFQ3 vs urine 3.
Abbreviations: FFQ, Food frequency questionnaire.

Table 7 Intraclass correlation coefficients for nutrients among the four food frequency questionnaire responses, by sex and overall

Nutrient	*Men*	*Women*	*Overall*
Energy	0.63	0.74	0.76
Carbohydrate	0.51	0.76	0.75
Protein	0.60	0.77	0.76
Fat	0.50	0.77	0.72
SFA	0.67	0.82	0.79
MUFA	0.44	0.74	0.69
PUFA	0.27	0.65	0.66
Cholesterol	0.84	0.88	0.87
Vitamin C	0.71	0.87	0.83
Retinol	0.88	0.90	0.89
Vitamin E	0.67	0.83	0.78
β-Carotene	0.76	0.87	0.84

Abbreviations: MUFA, mono unsaturated fatty acid; PUFA, poly unsaturated fatty acid; SFA, saturated fatty acid.

correlations were lower in men compared to women, particularly for fat and its components.

Discussion

In preparation for a prospective cohort study, we analyzed the validity and reliability of a newly designed FFQ in a population at high risk for esophageal cancer in northern Iran. The FFQ was based on the distinct cultural practices of the area, which is mainly populated by people of Turkmen and Persian ethnicity. Validity was assessed by comparing the results obtained from multiple FFQs with those from multiple 24-h diet recalls and biochemical markers of dietary intake in serum and urine. Reliability (reproducibility) was assessed by intraclass correlation coefficients between results of four FFQs administered to the same participants. The results showed good validity and reliability of the FFQ, with the notable exception of vitamin E.

The 24-h diet recalls were chosen as the reference method for assessment of validity because they were expected to have high response rate and good quality of response, and not to interfere much with the normal dietary habits of the subjects (Bohlscheid-Thomas *et al*., 1997a). USDA database was chosen in our study because of lack of any complete Iranian FCT (Azar and Sarkisian, 1981). An Iranian FCT was developed in 1980 for raw materials (Azar and Sarkisian, 1981) and it does not include the nutrients that come from cooked materials. Unpublished report shows that Iranian FCT about energy and macronutrients in bread as main Iranian food item is acceptable. The observed correlation coefficients between the FFQ and the 24-h recalls were high for all nutrients, except vitamin E, saturated fatty acid and MUFA. These results are similar to or better than those observed in similar studies (EPIC Group of Spain, 1997a, b; Bohlscheid-Thomas *et al*., 1997b; Kaaks *et al*., 1997; Katsouyanni *et al*., 1997), although some authors (Ocke *et al*., 1997b) reported higher correlations, particularly among men. Comparisons between questionnaire methods, although useful, are of limited use in gauging validity, because all these methods of dietary assessment are subject to reporting errors. To examine validity, comparisons were also made between the intake values obtained by either the FFQ or the 24-h recalls and the biochemical markers of dietary intake (Bingham *et al*., 1997). The latter has the advantage of offering quantitative measures unaffected by the subject's memory and the capacity to describe the food they consume. Frequently used biomarkers are blood levels of carotenoids and vitamin C for fruits and vegetables intake, whereas urinary nitrogen is a commonly employed and well-established marker of protein intake. In our study, two FFQs (FFQ1 and FFQ3) were administered at the time of blood collection, and one might expect a high correlation between serum biomarkers and intake from these FFQs. This might explain the better performance of FFQ3 – but not FFQ1 – as compared to the other FFQs and the average of 24-h recalls.

The analysis of reproducibility of plasma carotenoid levels might have been hampered by seasonal effect as blood samples were collected at 6 months intervals, at the end of summer and winter, and availability of fresh fruits in the study area varies by season. This can explain the lower correlation in our study as compared with a similar study from Europe, where importation of diverse fruits allows consumption throughout the year (Boeing *et al*., 1997).

The major dietary sources of vitamin C are orange and tangerine, which are mainly consumed during the cold months. It may explain the higher blood level of vitamin C in second measurement. We propose that in our population, the average value from several measurements of carotenoids in blood may be a better marker of long-term intake than a single measurement. Several intrinsic factors (metabolism, genetics) and extrinsic factors (physical activity, tobacco smoking), in addition to fruit and vegetables intake may influence the concentration of carotenoids in blood. It is important to control for these factors when carotenoids in blood are used as biological markers of the intake of fruit and vegetables.

We expected errors associated with fluctuations in the frequency of consumption of micronutrients such as vitamin C and carotenoids to be reduced by repeated observations with 24-h diet recall methods, as compared with the FFQ, which is designed to assess usual intake, as shown in studies reporting high correlations between serum biomarker level and dietary intake estimates based on 24-h recalls (Bingham *et al*., 1997). However, in our study, correlations between serum levels and estimates of intake of vitamin C, retinol and β-carotene, derived from a 24-h diet recalls were similar to those obtained from the FFQ.

The relationship between serum concentration of α-tocopherol and intake of vitamin E is confounded and attenuated by individual variations in absorption, availability and metabolism, and by the limited quality of food composition data for vitamin E (Ocke *et al*., 1997b). These notions, and the low correlations detected between serum α-tocopherol level and vitamin E intake cast doubts on whether this micronutrient can be assessed in a population-based study. Measuring creatinine in 24-h urine collections allowed us to assess the validity of the questionnaires for determining total protein intake (Bingham, 2003). Although the para amino benzoic acid method has also been proposed as more reliable (Bingham, 2003), we believe that our method is more practical because of the special ethnic and cultural situation among our subjects. In our study, the agreement between FFQ-based estimate of protein intake and 24-h urine nitrogen is similar to previous studies, but this agreement with 24-h dietary recall is lower than in some other studies (Bingham *et al*., 1997).

Overall, the reliability of the FFQs completed at 3-month intervals was good or moderately good (Bland and Altman, 1990). The reproducibility of FFQ has been examined under a wide variety of conditions, which complicates the comparison of results among different studies. Previous work reviewed by Willett (1990) has shown that the FFQ method gives reproducible nutrient intake estimates for middle-aged subjects, with correlations usually ranging from 0.5 to 0.8. More recent studies confirm these findings (Pietinen *et al*., 1988; Bueno-de-Mesquita *et al*., 1992; Rimm *et al*., 1992). Although the intraclass correlation analysis avoids the problem of a linear relationship being mistaken for agreement, its results are dependent on the range of the measurements (Bland and Altman, 1990). Therefore, the apparently high correlations detected in our study still suggest caution in the interpretation of results based on FFQ.

In conclusion, we have developed a population-specific FFQ for northeastern Iran. Using a combination of biochemical markers in urine and serum and repeated 24-h dietary recall, we have demonstrated that for most dietary factors of interest, the FFQ is sufficiently reliable and valid in this population of middle-aged subjects in a developing country.

Acknowledgements

This study was supported by funds from DDRC, Iran and IARC, France. We thank Dr Farhad Islami, Dr Hajiamin Marjani, Dr Omid Mozaffari, and Dr Elham Jafari for their logistic assistance. We sincerely thank our interviewers; Ailar Jamali, Anita Ramiar, Azin Kor, Akram Mohamadi, Afsane Mosadegh, Halime Pourghaz, Aijamal Gorgani, Fahime Igder, Saiideh Valaee, Homeira Davoodpour and Hesam Aman baee. We also enjoyed the close collaboration of Golestan health deputy Dr Mohamad Naemi, Dr Ali Aghapor, Mr J Kabir and Dr Jalil Pestei, Chief of Gonbad Health District.

We specially thank and appreciate Mrs Yoosefi and Mr Maramaie (local health workers of Behvarz) in the study areas. We also express appreciations to all the kind people of Gonbad for their collaboration in this study. We are deeply grateful to Elio Riboli and Nadia Slimani from IARC for their valuable comments.

References

Azar M, Sarkisian E (1981). *Food composition table of Iran*. National Nutrition and Food Science Research Institute of Shaheed Beheshti University.

Bachor PS, Denke MA, Stein EA (2001). Lipids and dyslipoproteinemia. In Henry JB (ed), *Clinical Diagnosis and Management by Laboratory Method* 20 edn. Saunders: London, pp 224–245.

Bingham SA (2003). Urine nitrogen as a biomarker for the validation of dietary protein intake. *J Nutr* **133**, 921S–924S.

Bingham SA, Gill C, Welch A, Cassidy A, Runswick SA, Oakes S *et al*. (1997). Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers. *Int J Epidemiol* **26** (Suppl 1), S137–S151.

Bland JM, Altman DG (1990). A note on the use of the intra-class correlation coefficient in the evaluation of agreement between two methods of measurement. *Comput Biol Med* **30**, 337–340.

Boeing H, Bohlscheid-Thomas S, Voss S, Schneewiss S, Wahrendoorf J (1997). The relative validity of vitamin intakes derived from a food frequency questionnaire compared to 24-h recalls and biological measurements: results from the EPIC Pilot Study in Germany. *Int J Epidemiol* **26** (Suppl 1), S82–S90.

Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J (1997a). Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. *Int J Epidemiol* **26** (Suppl 1), S59–S70.

Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J (1997b). Reproducibility and relative validity of energy and macronutrient

intake of a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26** (Suppl 1), S71–S81.

Bueno-de-Mesquita HB, Smeets FW, Runia S, Hulshof KFAM (1992). The reproducibility of a food frequency questionnaire among controls participating in a case-control study on cancer. *Nutr Cancer* **18**, 143–156.

Cook-Mozaffari PJ, Azordegan F, Day NE, Ressicaud A, Sabai C, Armesh B (1979). Esophageal cancer studies in the Caspian littoral of Iran: results of a case-control study. *Br J Cancer* **39**, 293–310.

EPIC Group of Spain (1997a). Relative validity and reproducibility of a diet history questionnaire in Spain. II. Nutrients. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26** (Suppl 1), S100–S109.

EPIC Group of Spain (1997b). Relative validity and reproducibility of a diet history questionnaire in Spain. III. Biochemical markers. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26** (Suppl 1), S110–S117.

EPIC Group of Spain (1997c). Relative validity and reproducibility of a diet history questionnaire in Spain. I. Foods. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26** (Suppl 1), S91–S99.

Hormozdiari H, Day NE, Aramesh B, Mahboubi E (1975). Dietary factors and esophageal cancer in the Caspian Littoral of Iran. *Cancer Res* **35**, 3493–3498.

Hristova EN, Henry JB (2001). Metabolic intermediats. In organic ions and biochemical markers of bone metabolism. In Henry JB (ed), *Clinical Diagnosis and Management by Laboratory Method*, 20 edn. Saunders: London, pp 180–209.

Hunter D (1990). Biochemical indicators of dietary intake. In: Willett W (ed). *Nutritional Epidemiology*. Oxford University Press: New York, pp. 143–216.

Isaksson B (1980). Urinary nitrogen output as a validity test in dietary surveys. *Am J Clin Nutr* **33**, 4–5.

Islami F, Kamangar F, Aghcheli K, Fahimi S, Semnani S, Taghavi N *et al*. (2004). Epidemiologic features of upper gastrointestinal tract cancers in Northeastern Iran. *Br J Cancer* **90**, 1402–1406.

Kaaks R, Slimani N, Riboli E (1997). Pilot phase studies on the accuracy of dietary intake measurements in the EPIC project: overall evaluation of results. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26** (Suppl 1), S26–S36.

Katsouyanni K, Rimm EB, Gnardellis C, Trichopoulos D, Polychronopulos E, Trichopoulou A (1997). Reproducibility and relative validity of an extensive semi-quantitative food frequency questionnaire using Dietary records and biochemical markers among Greek Schoolteachers. *Int J Epidemiol* **26** (Suppl 1), S118–S127.

Kmet J, Mahboubi E (1972). Esophageal cancer in the Caspian littoral of Iran: initial studies. *Science* **175**, 846–853.

Kroke A, Klipstein-Grobusch K, Voss S, Moseneder J, Thielecke F, Noack R *et al*. (1999a). Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *Am J Clin Nutr* **70**, 439–447.

Lucas F, Niravong M, Villeminot S, Kaaks R, Lavel-Chapelon F (1995). Estimation of food portion size using photograph: validity, strengths, weaknesses and recomendations. *J Hum Nutri Dietetics* **8**, 65–74.

Ocke MC, Bueno-de-Mesquita HB, Goddijn HE, Jansen A, Pols MA, Van Staveren WA *et al*. (1997a). The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol* **26** (Suppl 1), S37–S48.

Ocke MC, Bueno-de-Mesquita HB, Pols MA, Smit HA, Van Staveren WA, Kromhout D (1997b). The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *Int J Epidemiol* **26** (Suppl 1), S49–S58.

Pietinen P, Hartman AM, Haapa E, Pasanen L, Haapaksoki J, Palmgren J *et al*. (1988). Reproducibility and validity of dietary assessment instruments I. A self-administered food use questionnaire with a portion size picture booklet. *Am J Epidemiol* **128**, 655–666.

Pourshams A, Saadatian-Elahi M, Nouraie M, Malekshah AF, Rakhshani N, Salahi R *et al*. (2005). Golestan cohort study of oesophageal cancer: feasibility and first results. *Br J Cancer* **92**, 176–181.

Prentice RL, Sugar E, Wang CY, Neuhouser M, Patterson R (2002). Research strategies and the use of nutrient biomarkers in studies of diet and chronic disease. *Public Health Nutr* **5**, 977–984.

Rimm EJ, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC (1992). Reproducibility and validity of an expanded self-administered semi quantitative food frequency questionnaire among male health professionals. *Am J Epidemiol* **135**, 1114–1126.

Steghens JP, van Kappel AL, Riboli E, Collombel C (1997). Simultaneous measurement of seven carotenoids, retinol and alpha-tocopherol in serum by high-performance liquid chromatography. *J Chromatogr B Biomed Sci Appl* **20**, 71–81.

Vuilleumier JP, Keck E (1989). Fluorometric assay of vitamin C in biological materials using a centrifugal analyser with fluorescence attachment. *J Micronutr Anal* **5**, 25–34.

Willett WC (1990). *Nutrition Epidemiology*. Oxford University Press: New York.